PCT/NZ00/00031

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

Intellectual Property Office of New Zealand
Te Pou Rāhui Hanga Hou

CERTIFICATE

This certificate is issued in support of an application for Patent registration in a country outside New Zealand pursuant to the Patents Act 1953 and the Regulations thereunder.

I hereby certify that annexed is a true copy of the Provisional Specification as filed on 17 March 1999 with an application for Letters Patent number 334715 made by MICHIGAN TECHNOLOGICAL UNIVERSITY;CARTER HOLT HARVEY LTD;FLETCHER CHALLENGE FORESTS LTD;DAVID F KARNOSKY;GOPI KRISHNA PODILA;JUN JUN LIU.

Dated 21 March 2000.

Neville Harris
Commissioner of Patents

334715

Patents Form No. 4

PATENTS ACT 1953

PROVISIONAL SPECIFICATION

PLANTS HAVING MODIFIED REPRODUCTIVE CAPACITY

We, **GOPI KRISHNA PODILA**, an Indian citizen of 307 West 42nd Street, Houghton, MI 49931, United States of America, **JUN-JUN LIU**, a Chinese citizen of #7, 900 Champion Street, Houghton, MI 49931, United States of America, and **DAVID F KARNOSKY,** a US citizen of Route 1, Box 139, Chassell, Michigan, United States of America, in trust for **CARTER HOLT HARVEY LIMITED**, a company duly incorporated under the laws of New Zealand of 640 Great South Road, Manakau City, Auckland, New Zealand, **FLETCHER CHALLENGE FORESTS LIMITED**, a New Zealand company of 3 Rockridge Avenue, Penrose, Auckland, New Zealand, and **MICHIGAN TECHNOLOGICAL UNIVERSITY** of 1400 Townsend Drive, Houghton, Michigan 49931, United States of America, do hereby declare this invention to be described in the following statement:

(followed by page 1A)

PLANTS HAVING MODIFIED REPRODUCTIVE CAPACITY

This invention relates to plants having modified reproductive capacity. In particular, it relates to woody plants of the *Pinus* genus which have been genetically modified to have diminished reproductive capacity or which are sterile.

INTRODUCTION

It is desirable that genes be identified which are involved in reproduction/flower development in plants and in particular in woody plants of the *Pinus* genus. Knowledge of such genes allows for modulation of the reproductive capacity of *Pinus* plants and specifically the production of reproductively null (sterile) plants.

The applicants have now identified and isolated such a reproductive gene which encodes a peptide involved in the reproductive cycle of *Pinus radiata*. It is broadly towards this gene, to its homologs in other plants of the *Pinus* genus and to the modulation of its expression/function within plants of the *Pinus* genus that the present invention is directed.

The invention further provides the endogenous reproductive-tissue-specific promoter for the reproductive gene.

SUMMARY OF THE INVENTION

In a first aspect, the present invention broadly provides a plant of the *Pinus* genus, which contains at least a polynucleotide encoding a peptide having the amino acid sequence of Figure 1 or a functionally equivalent variant thereof, and which has been genetically modified to have at least a reduced reproductive capacity.

Preferably, said plant is sterile.

Conveniently, the plant is a member of the species *Pinus radiata, Pinus taeda, Pinus elliotti, Pinus clausa, Pinus palustrus, Pinus echinata, Pinus ponderosa, Pinus jeffrey, Pinus resinosa, Pinus rigida, Pinus banksiana, Pinus serotina, Pinus strobus, Pinus monticola, Pinus lambertiana, Pinus virginiana, Pinus contorta, Pinus cariboea, Pinus*

pinaster, Pinus brutia, Pinus eldarica, Pinus coulteri, Pinus nigra, Pinus sylvestris, Pinus tecunumannii, Pinus keysia, Pinus oocarpa and *Pinus maxinumoii.*

Preferably, the genetic modification reduces the functional expression of said peptide (ie. the amount of the peptide which is expressed and functional within the plant).

In a further embodiment, the invention provides a polynucleotide which encodes a peptide having the amino acid sequence of Figure 1 or a variant thereof.

Most preferably, said polynucleotide includes part or all of the nucleotide sequence of Figure 1.

Preferably, the polynucleotide is DNA.

The invention further provides a DNA construct which includes a polynucleotide as defined above.

More particularly, the invention provides a DNA construct comprising, in the 5'-3' direction:

(a) a promoter sequence;
(b) an open reading frame polynucleotide coding for the peptide having the amino acid sequence of Figure 1 or a functionally equivalent variant thereof; and
(c) a termination sequence.

In one embodiment, the open reading frame is in a sense orientation.

In an alternative embodiment, the open reading frame is in an anti-sense orientation.

In still a further embodiment, the invention provides a DNA construct comprising, in the 5'-3' direction:

(a) a promoter sequence;

(b) a non-coding region of a gene coding for the peptide having the amino acid sequence of Figure 1 or a functionally equivalent variant thereof; and

(c) a termination sequence.

Once again, the non-coding region can be in a sense or anti-sense orientation.

In yet a further embodiment, the invention provides a DNA construct comprising, in the 5'-3' direction:

(a) a promoter sequence;

(b) a polynucleotide comprising a nucleotide sequence complementary to at least part of a sequence coding for the peptide having the amino acid sequence of Figure 1 or a functionally equivalent variant thereof; and

(c) a termination sequence.

In still a further embodiment, the invention provides a DNA construct comprising, in the 5'-3' direction:

(a) a promoter sequence;

(b) a polynucleotide comprising a nucleotide sequence encoding the DNA binding domain of the peptide having the amino acid sequence of Figure 1 or a functionally equivalent variant thereof, but not the remainder of the peptide of Figure 1; and

(c) a termination sequence.

Preferably, in each embodiment, the construct further includes a marker for identification of transformed cells.

Preferably, in each construct the promoter sequence is, or is based upon, the sequence of Figure 2.

In yet a further embodiment, the invention provides a reproductive-tissue-specific promoter which has the nucleotide sequence of Figure 2, or which has a nucleotide sequence which is a functionally equivalent variant thereof.

The invention further provides a DNA construct which includes a promoter as defined above.

More specifically, said construct comprises, in the 5'-3' direction:

(a) the reproductive-tissue-specific promoter as defined above;

(b) an open reading frame polynucleotide encoding a peptide which is to be expressed in plant reproductive tissue; and

(c) a termination sequence.

Polynucleotide (b) may encode a peptide which at least reduces the functional capacity of said plant reproductive tissue, preferably such that a plant transformed with said construct is rendered sterile.

In still a further aspect, the invention provides a transgenic plant cell which includes a DNA construct as defined above.

DESCRIPTION OF THE DRAWINGS

While the invention is broadly as defined above, it will be appreciated by those persons skilled in the art that it is not limited thereto and that it further includes embodiments of which the following description provides examples. In addition, the invention will be better understood through reference to the accompanying drawings in which:

Figure 1 shows the amino acid sequence of the reproductive peptide of the invention, PrAG1, together with the nucleotide sequence coding therefor;

Figure 2 shows the sequence of the endogenous promoter for PrAG1, isolated from *Pinus radiata*;

Figure 3 is an RNA gel blot analysis of PrAG1 mRNA accumulation in *Pinus radiata* organs. Twenty μg of total RNA from various organs was electrophoresed, blotted onto nylon membranes, and hybridized with 3'-terminal fragment of pRAG1 cDNA. Total RNA was isolated from immature male cone(M), immature female cone(F),

vegetable shoot (V)s, needle (N) and stem (S). The 26S and 18S rRNA was used as control (bottom);

Figure 4 is a DNA gel blot analysis of *Pinus radiata* genomic DNA hybridized with the 3' terminal region of PrAG1. 20 μ genomic DNA was digested with BamHI (BA) and Bgl II (BG) EcoRI(E), HindIII (H), XhoI(X); and

Figure 5 is a Reverse Transcription-Polymerase Chain Reaction (RT-PCR) analysis showing reproductive-organ specific expression of PrAG1. RT-PCR analysis was performed on total RNA isolated from different organs of radiata Pine: (M) immature male cone, (F) immature female cone, (Vs) vegetative shoot, (N) needle and (S) stem. MADS box genes were amplified with PrAG1 gene gene-specific oligonucleotides. Products from the PCR reactions were electrophoresed, blotted, and hybridized with a labelled probe of PrAG1 specific fragment.

DESCRIPTION OF THE INVENTION

As broadly outlined above, the applicants have identified a peptide which is involved in plant reproduction, together with the gene coding therefor. The specific peptide and gene are from *Pinus radiata*.

The amino acid sequence of the peptide and its encoding nucleotide sequence are given in Figure 1. It will however be appreciated that the invention is not restricted only to the peptide/polynucleotide having the specific amino acid/nucleotide sequence given in Figure 1. Instead, the invention also extends to functionally equivalent variants of the peptide/polynucleotide of Figure 1.

The term "polynucleotide(s)" as used herein means a single or double-stranded polymer of deoxyribonucleotide or ribonucleotide bases and includes DNA and corresponding RNA molecules, including HnRNA and mRNA molecules, both sense and anti-sense strands, and comprehends cDNA, genomic DNA and recombinant DNA, as well as wholly or partially synthesized polynucleotides. An HnRNA molecule contains introns and corresponds to a DNA molecule in a generally one-to-one manner. An mRNA molecule corresponds to an HnRNA and DNA molecule from which the introns have been excised. A polynucleotide may consist of an entire gene, or any portion thereof. Operable anti-sense polynucleotides may comprise a

fragment of the corresponding polynucleotide, and the definition of "polynucleotide" therefore includes all such operable anti-sense fragments.

The phrase "functionally equivalent variants" recognises that it is possible to vary the amino acid/nucleotide sequence of a peptide while retaining substantially equivalent functionality. For example, a peptide can be considered a functional equivalent of another peptide for a specific function if the equivalent peptide is immunologically cross-reactive with and has at least substantially the same function as the original peptide. The equivalent can be, for example, a fragment of the peptide, a fusion of the peptide with another peptide or carrier, or a fusion of a fragment which additional amino acids. For example, it is possible to substitute amino acids in a sequence with equivalent amino acids using conventional techniques. Groups of amino acids normally held to be equivalent are:

(a) Ala, Ser, Thr, Pro, Gly;
(b) Asn, Asp, Glu, Gln;
(c) His, Arg, Lys;
(d) Met, Leu, Ile, Val; and
(e) Phe, Tyr, Trp.

Equally, nucleotide sequences encoding a particular product can vary significantly simply due to the degeneracy of the nucleic acid code.

Variants can have a greater or lesser degree of homology as between the variant amino acid/nucleotide sequence and the original.

Polynucleotide or polypeptide sequences may be aligned, and percentage of identical nucleotides in a specified region may be determined against another sequence, using computer algorithms that are publicly available. Two exemplary algorithms for aligning and identifying the similarity of polynucleotide sequences are the BLASTN and FASTA algorithms. The similarity of polypeptide sequences may be examined using the BLASTP algorithm. Both the BLASTN and BLASTP software are available on the NCBI anonymous FTP server (ftp://ncbi.nlm.nih.gov) under /blast/executables/. The BLASTN algorithm version 2.0.4 [Feb-24-1998], set to the default parameters described in the documentation of variants according to the present invention. The use of the BLAST family of algorithms, including BLASTN

and BLASTP, is described at NCBI's website at URL http://www.ncbi.nlm.nih.gov/BLAST/newblast.html and in the publication of Altschul, Stephen F., *et al.* (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", *Nucleic Acids Res.* 25:3389-34023. The computer algorithm FASTA is available on the Internet at the ftp site ftp://ftp.virginia.edu/pub/fasta/. Version 2.0u4, February 1996, set to the default parameters described in the documentation and distributed with the algorithm, is also preferred for use in the determination of variants according to the present invention. The use of the FASTA algorithm is described in W. R. Pearson and D. J. Lipman, "Improved Tools for Biological Sequence Analysis", *Proc. Natl. Acad. Sci. USA* 85:2444-2448 (1988) and W. R. Pearson, "Rapid and Sensitive Sequence Comparison with FASTP and FASTA, "*Methods in Enzymology* 183:63-98 (1990).

The following running parameters are preferred for determination of alignments and similarities using BLASTN that contribute to E values (as discussed below) and percentage identity: Unix running command: blastall -p blastn -d embldb -e 10 -G 1 -E 1 -r 2 -v 50 -b 50 -I queryseq -o results; and parameter default values:

-p Program Name [String]

-d Database [String]

-e Expectation value (E) [Real]

-G Cost to open a gap (zero invokes default behaviour) [Integer]

-E Cost to extend a cap (zero invokes default behaviour) [Integer]

-r Reward for a nucleotide match (blastn only) [Integer]

-v Number of one-line descriptions (V) [Integer]

-b Number of alignments to show (B) [Integer]

-i Query File [File In]

-o BLAST report Output File [File Out] Optional

For BLASTP the following running parameters are preferred: blastall -p blastp -d swissprotdb -e 10 -G 1 -E 1 -v 50 -b 50 -I queryseq -o results

-p Program Name [String]

-d Database [String]

-e Expectation value (E) [Real]

-G Cost to open a gap (zero invokes default behaviour) [Integer]

-E Cost to extend a cap (zero invokes default behaviour) [Integer]

-v Number of one-line descriptions (v) [Integer]

-b Number of alignments to show (b) [Integer]

-i Query File [File In]

-o BLAST report Output File [File Out] Optional

The "hits" to one or more database sequences by a queried sequence produced by BLASTN, BLASTP, FASTA, or a similar algorithm, align and identify similar portions of sequences. The hits are arranged in order of the degree of similarity and the length of sequence overlap. Hits to a database sequence generally represent an overlap over only a fraction of the sequence length of the queried sequence.

The BLASTN and FASTA algorithms also produce "Expect" or E values for alignments. The E value indicates the number of hits one can "expect" to see over a certain number of contiguous sequences by chance when searching a database of a certain size. The Expect value is used as a significance threshold for determining whether the hit to a database, such as the preferred EMBL database, indicates true similarity. For example, an E value of 0.1 assigned to a hit is interpreted as meaning that in a database of the size of the EMBL database, one might expect to see 0.1 matches over the aligned portion of the sequence with a similar score simply by chance. By this criterion, the aligned and matched portions of the sequences then have a 90% probability of being the same. For sequences having an E value of 0.01 or less over aligned and matched portions, the probability of finding a match by chance in the EMBL database is 1% or less using the BLASTN or FASTA algorithm.

According to one embodiment, "variant" polynucleotides, with reference to each of the polynucleotides of the present invention, preferably comprise sequences having the same number or fewer nucleic acids than each of the polynucleotides of the present invention and producing an E value of 0.01 or less when compared to the polynucleotide of the present invention. That is, a variant polynucleotide is any sequence that has at least a 99% probability of being the same as the polynucleotide of the present invention, measured as having an E value of 0.01 or less using the BLASTN or FASTA algorithms set at the parameters discussed above.

Variant polynucleotide sequences will generally hybridize to the recited polynucleotide sequence under stringent conditions. As used herein, "stringent conditions" refers to prewashing in a solution of 6X SSC, 0.2% SDS; hybridizing at 65°C, 6X SSC, 0.2% SDS overnight; followed by two washes of 30 minutes each in

1X SSC, 0.1% SDS at 65°C and two washes of 30 minutes each in 0.2X SSC, 0.1% SDS at 65°C.

It is of course expressly contemplated that homologs to PrAG1 exist in other members of the *Pinus* genus. Such homologs are also "functionally equivalent variants" of PrAG1 as the phrase is used herein.

DNA sequences from *Pinus* plants other than *Pinus radiata* which are homologs of PrAG1 may be isolated by high throughput sequencing of cDNA libraries prepared from such *Pinus* plants. Alternatively, oligonucleotide probes based on the sequences for PrAG1 provided in Figure 1 can be synthesized and used to identify positive clones in either cDNA or genomic DNA libraries from other *Pinus* plants by means of hybridization or PCR techniques. Probes should be at least about 10, preferably at least about 15 and most preferably at least about 20 nucleotides in length. Hybridization and PCR techniques suitable for use with such oligonucleotide probes are well known in the art. Positive clones may be analyzed by restriction enzyme digestion, DNA sequencing or the like.

The polynucleotides of the present invention may be generated by synthetic means using techniques well known in the art. Equipment for automated synthesis of oligonucleotides is commercially available from suppliers such as Perkin Elmer/Applied Biosystems Division (Foster City, CA) and may be operated according to the manufacturer's instructions.

The primary importance of identification of the reproductive peptide/polynucleotides of the invention is that they enable the reproductive capacity of *Pinus* plants to be modulated. This modulation will generally involve a reduction in the functional expression (silencing) of the reproductive peptide.

Any conventional technique for effecting this can be employed. Examples include co-suppression or anti-sense strategies, a dominant negative approach, or techniques which involve ribozymes to digest, or otherwise be lethal to, RNA post-transcription of the target gene.

Co-suppression can be effected in a manner similar to that discussed, for example, by Napoli *et al* (Plant Cell 2:279-290, 1990) and de Carvalho Niebel *et al* (Plant Cell

7:347-258, 1995). In some cases, it can involve overexpression of the gene of interest through use of a constitutive promoter. It can also involve transformation of a plant with a non-coding region of the gene, such as an intron from the gene or 5'-non-coding leader sequences.

Anti-sense strategies involve expression or transcription of DNA with the expression/transcription product being capable of interfering with translation of mRNA transcribed from the target gene. This will normally be through the expression/transcription product hybridising to and forming a duplex with the target mRNA.

The expression/transcription product can be a relatively small molecule and still be capable of disrupting mRNA translation. However, the same result is achieved by expressing the target gene in an anti-sense orientation such that the RNA produced by transcription of the anti-sense oriented gene is complementary to all or part of the endogenous target mRNA.

Anti-sense strategies are described generally by Robinson-Benion *et al.*, (1995), Anti-sense techniques, *Methods in Enzymol.* 254(23):363-375 and Kawasaki *et al.*, (1996), *Artific. Organs* 20 (8): 836-848.

Dominant negative approaches involve the expression of a modified DNA binding/activating protein which includes a DNA binding domain but not a activator domain. The result is that the protein binds to DNA as intended but fails to activate, while at the same time blocking the binding of the DNA binding/activating peptides which normally bind to the same site.

The ribozyme approach to regulation of peptide expression involves inserting appropriate sequences or subsequences (eg. DNA or RNA) in ribozyme constructs (McIntyre CL, Manners JM, *Transgenic Res.* 5(4):257-262, 1996). Ribozymes are synthetic RNA molecules that comprise a hybridizing region complementary to two regions, each of which comprises at least 5 contiguous nucleotides of a mRNA molecule encoded by one of the inventive polynucleotides. Ribozymes possess highly specific endonuclease activity, which autocatalytically cleaves the mRNA.

To give effect to the above strategies, the invention also provides DNA constructs. The constructs include the intended DNA (such as the gene of the invention in anti-sense orientation or a polynucleotide encoding the appropriate DNA binding domain or ribozyme), a promoter sequence and a termination sequence, operably linked to the DNA sequence to be transcribed, which control expression of the gene. The promoter sequence is generally positioned at the 5' end of the DNA sequence to be transcribed, and is employed to initiate transcription of the DNA sequence. Promoter sequences are generally found in the 5' non-coding region of a gene but they may exist in introns (Luehrsen, K.R., *Mol. Gen. Genet.* 225:81-93, 1991) or in the coding region. When the construct includes an open reading frame in a sense orientation (for co-suppression through over-expression) the promoter sequence also initiates translation of the open reading frame. For DNA constructs comprising either an open reading frame in an anti-sense orientation or a non-coding region, the promoter sequence generally consists only of a transcription initiation site having a RNA polymerase binding site.

A variety of promoter sequences which may be usefully employed in the DNA constructs of the present invention are well known in the art. The promoter sequence, and also the termination sequence, may be endogenous to the target *Pinus* plant host or may be exogenous, provided the promoter is functional in the target host. For example, the promoter and termination sequences may be from other plant species, plant viruses, bacterial plasmids and the like. Preferably, promoter and termination sequences are those endogenously associated with the reproductive genes.

The structure and sequence of the endogenous promoter for PrAG1 (which is from *Pinus* radiata) is shown in Figure 2.

Factors influencing the choice of promoter include the desired tissue specificity of the construct, and the timing of transcription and translation. For example, constitutive promoters, such as the 35S Cauliflower Mosaic Virus (CaMV 35S) promoter, will affect the activity in all parts of the plant. Use of a tissue specific promoter will result in production of the desired sense or antisense RNA only in the tissue of interest. With DNA constructs employing inducible promoter sequences, the rate of RNA polymerase binding and initiation can be modulated by external stimuli, such as light, heat, anaerobic stress, alteration in nutrient conditions and

the like. Temporally regulated promoters can be employed to effect modulation of the rate of RNA polymerase binding and initiation at a specific time during development of a transformed cell. Preferably, the original promoters from the gene in question, or promoters from a specific tissue-targeted gene in the organism to be transformed are used. Other examples of promoters which may be usefully employed in the present invention include, mannopine synthase (mas), octopine synthase (ocs) and those reviewed by Chua *et al.* (Science, 244:174-181, 1989).

The termination sequence, which is located 3' to the DNA sequence to be transcribed, may come from the same gene as the promoter sequence or may be from a different gene. Many termination sequences known in the art may be usefully employed in the present invention, such as the 3' end of the *Agrobacterium tumefaciens* nopaline synthase gene. However, preferred termination sequences are those from the original gene or from the target *Pinus* species to be transformed.

The DNA constructs of the present invention may also contain a selection marker that is effective in *Pinus* plant cells, to allow for the detection of transformed cells containing the construct. Such markers, which are well known in the art, typically confer resistance to one or more toxins. One example of such a marker is the NPTII gene whose expression results in resistance to kanamycin or hygromycin, antibiotics which is usually toxic to plant cells at a moderate concentration (Rogers *et al.*, in *Methods for Plant Molecular Biology*, A Weissbach and H Weissbach eds, Academic Press Inc., San Diego, CA (1988)). Alternatively, the presence of the desired construct in transformed cells can be determined by means of other techniques well known in the art, such as Southern and Western blots.

Techniques for operatively linking the components of the inventive DNA constructs are well known in the art and include the use of synthetic linkers containing one or more restriction endonuclease sites as described, for example, by Maniatis *et al.*, (*Molecular Cloning: A Laboratory Manual*, Cold Spring Harbour Laboratories, Cold Spring Harbour, NY, 1989). The DNA construct may be linked to a vector having at least one replication system, for example, *E. coli*, whereby after each manipulation, the resulting construct can be cloned and sequenced and the correctness of the manipulation determined.

The DNA constructs of the present invention may be used to transform a variety of plants of the *Pinus* genus. In a preferred embodiment, the DNA constructs are employed to transform *Pinus radiata, Pinus taeda, Pinus elliotti, Pinus clausa, Pinus palustrus, Pinus echinata, Pinus ponderosa, Pinus jeffrey, Pinus resinosa, Pinus rigida, Pinus banksiana, Pinus serotina, Pinus strobus, Pinus monticola, Pinus lambertiana, Pinus virginiana, Pinus contorta, Pinus cariboea, Pinus pinaster, Pinus brutia, Pinus eldarica, Pinus coulteri, Pinus nigra, Pinus sylvestris, Pinus tecunumannii, Pinus keysia, Pinus oocarpa* and *Pinus maxinumoii.*

As discussed above, transformation of a plant of the *Pinus* genus with a DNA construct including an open reading frame coding for a peptide encoded by a DNA sequence of the invention wherein the open reading frame is orientated in a sense direction can, in some cases, lead to a decrease in expression of the peptide by co-suppression. Transformation of the plant with a DNA construct comprising an open reading frame in an anti-sense orientation or a non-coding (untranslated) region of a gene will lead to a decrease in the expression of the peptide in the transformed plant.

Techniques for stably incorporating DNA constructs into the genome of target plants are well known in the art and include *Agrobacterium tumefaciens* mediated introduction, electroporation, protoplast fusion, injection into reproductive organs, injection into immature embryos, high velocity projectile introduction and the like. The choice of technique will depend upon the target plant to be transformed.

Once the cells are transformed, cells having the DNA construct incorporated into their genome may be selected by means of a marker, such as the kanamycin resistance marker discussed above. Transgenic cells may then be cultured in an appropriate medium to regenerate whole plants, using techniques well known in the art. In the case of protoplasts, the cell wall is allowed to reform under appropriate osmotic conditions. In the case of seeds or embryos, an appropriate germination or callus initation medium is employed. For explants, an appropriate regeneration medium is used.

For a review of regeneration of forest trees such as those of the *Pinus* genus, see Dunstan *et al.,* Somatic embryogenesis in woody plants. In: Thorpe, T.A. ed. 1995:

in vitro embryogenesis of plants. Vol 20 in Current Plant Science and Biotechnology in Agriculture, Chapter 12, pp. 471-540.

The resulting transformed plants may be reproduced sexually or asexually, using methods well known in the art, to give successive generations of transgenic plants.

The endogenous promoter for PrAG1 forms a further, but important, aspect of the invention. As mentioned above, the promoter has the nucleotide sequence of Figure 2 although, again, functionally equivalent variants are not excluded.

The endogenous promoter is reproductive-tissue-specific. This means that it can be employed in constructs to express any desired peptide in plant reproductive tissue. This includes the peptide encoded by PrAG1, but can also be another peptide. That other peptide can be a peptide which, when produced, causes the reproductive organs of the plant to abort, redefine themselves as vegetative or stop development. The peptide encoded can, for example, also be a peptide causing cell death. Illustrative peptides/genes are Diphtheria Toxin A (DTA), Barnase (from Bacillus amyloliquefaciens), apoptosis genes, glucanase, and RNAses.

Alternatively, the peptide can be one which, when produced, alters the timing of flowering (ie. either delays or accelerates flowering, such as the ELF-3 and CONSTANS flowering time genes).

The peptide to be expressed can be ligated to the promoter in a sense or antisense orientation, dependant upon the desired effect.

The construct will also include a termination sequence. As above, this sequence will be selected depending upon the plant in which it is to be active.

Other optional elements can be included in the construct as discussed previously, including marker sequences, if these are desirable.

The promoter, and constructs containing it, are not restricted in use to plants of the *Pinus* genus. Instead, they can be used to transform other agronomically important plants in which modulation of reproductive capacity (particularly the timing and

abundance of flowering) is desirable. Such plants include cereals, rice, maise, wheat, barley, oats, rye, soyabean and canola.

The invention will now be described with reference to the following non-limiting examples.

EXPERIMENTAL

1. PrAG1 cDNA Cloning:

Total RNA was purified from immature cone of radiata pine according to the protocol of Charles Ainsworth (Plant Molecular Biology Reporter, 12(3), 1994: 198-203). The mRNA was isolated with oligo-T cellulose column. With mRNA as template, the cDNA was synthesised with Cap-Finder cDNA synthesis kit (ClonTech Co.). The cDNA was inserted into Lambda TriplEx Vector (ClonTech Co.), then packaged it with Gigapack@ III packaging extracts (Stratagene Co.) to obtain a cDNA library.

Two degenerate primers were designed:

3' PCR primer: 5' GCIGTIAGIYCITCICCCAT3' ;
5' PCR primer: 5' AAYCGICARGTIACITT3'

These primers were used to perform RT-PCR based differential screening on RNA from various female tissue sources, including immature female buds, vegetative tissue from needles and later stages of development. The 50-ul reaction mixture contained 2.5 Units Taq DNA polymerase, 1X Polymerization Buffer (both from ClonTech Co.), 1mMMgCl_2 , 0.2mMdNTP and 0.25uM primers. The PCR was performed under the following conditions: denaturation at 94°C for 30 s, annealing at 50°C for 1 min and extension at 72°C for 1 min for 30 cycles on a Thermal Cycler 480 (Perkin-Elmer, Norwalk, CT, USA).

Fragments were obtained mostly from immature female bud tissue RNA samples. Several DNA fragments were cloned into pGEM-T vector and sequenced. Sequence analysis showed that most of these fragments contained similar sequences. One of the DNA fragments (309 bp) was chosen. This was used as a probe to screen the cDNA library to clone its full-length cDNA and resulted in the cDNA clone, PrAG1.

The sequence of PrAG1 was analysed on both strands by the Sanger's dideoxy method (Sanger *et al.*, 1977, Proc. Natal. Acad. Sci. U.S.A. 74: using a Sequenase kit (United States Biochemical co.).

The resulting sequence is shown in Figure 1 gives the nucleotide sequence coding for the peptide of the invention together with the predicted amino acid sequence.

Sequence comparison and phylogenetic analysis were conducted with the software program MacDNASIS (Version 3.5, Hitachi Corp.). The results of analysis revealed PrAG1 to be a MADS box gene.

2. PrAG1 Promoter Cloning:

I. Genomic DNA purification:

Genomic DNA was purified from young needles according to a CTAB method as described below.

1) 2 g of young needles of *Pinus* radiata were ground in liquid nitrogen (mortar and pestle) to a fine powder.

2) This powder was mixed with 15 ml of pre-warmed CTAB extraction buffer [3% CTAB(W/V), 100mM Tris-HCl pH8.0, 20 mM EDTA pH8.0, 1.4 M NaCl, 1% PVP 940,000, 1% beta mercaptoethanol] and incubated at 65 C for one hour.

3) To the above mixture 15 mL chloroform was added and mixed gently.

4) The contents were centrifuged at 10,000g for 20 minutes at 4 C.

5) The supernatant was transferred to a new tube, and mixed with 1/10 volume of 3M sodium acetate (pH4.8), and 0.7 volume of isopropanol. The DNA was precipitated at -20 C for 30 minutes.

6) The DNA was pelleted at 10,000 g for 10 minutes at 4 C.

7) The DNA pellet was then air dried and resuspended in 2 mL TE buffer (10 mM Tris-HCl pH7.5, 1mM EDTA pH8.0) and 2 ul of RNAse A (10ug/uL) was added. The contents were incubated at 37 C for 30 minutes to remove any RNA from the sample.

8) After the incubation, 2 mL of 5M Ammonium acetate and 10 mL of 100% ethanol were added and the contents kept at -20 C for 15 minutes.

9) The mixture was then centrifuged at 10,000 g for 10 minutes at 4°C to pellet DNA. The DNA pellet was washed in 70% ethanol twice.

10) The DNA pellet was air dried and resuspended in 200 uL TE bufffer.

II. Cloning of PrAG1 Promoter with Two step Genomic DNA Walking.

1) The Universal Genome Walker Kit (CLONTECH) was employed. For the first step genomic DNA walking, two PrAG1 specific primers were designed and synthesized according to the PrAG1 cDNA sequence. The sequences of the primers were:

Primer GSP1: 5' CGC CTT CTT CAA TAA ACC ATT TCG GCG CTT 3'
Primer GSP2: 5' GAC CTG TCG GTT CGT AGT ATT TTC AAT CCT 3'

2) Based upon the promoter sequence we got from step 1), two PrAG1 promoter sequence specific primer were designed and synthesized. The primers were:

Primer GSP3: 5' TTC GTC CTC CAT TTT GTG CGC TCT CCA TTC 3'
Primer GSP4: 5' GCA CTC CAC TCT TCC TTT ATT TCT TAC CAC 3'.

3) According to the User Manual of Universal Genome Walker Kit, 13 genome walker libraries were constructed after genomic DNA digestion with restriction enzymes: EcoR V, Sca I, Dra I, Pvu II, Ssp I, Stu I, Sma I, Hap I, BsaB I, Bcl136 II, Pml I, Nru I, Hic II.

4) With 13 genome walker libraries as templates, and adaptor primer 1 (AP1 primer from kit) and GSP1 primer, first round PCR was performed under the conditions suggested by the kit manufacturer. After agarose electrophoresis analysis of the

PCR product, second round PCR was performed with the nested primers AP2 (Adaptor primer from the kit) and GSP2. The PCR products from the second round PCR were purified and cloned into pGEM-T easy vector (Promega). Following sequence analysis, and DNA sequence comparison with PrAG1 cDNA, one DNA fragment of 1105 bp from Sca I genome walker library was obtained which was identified as the promoter region of PrAG1, based upon the overlapped region between it and PrAG1 cDNA.

5) The second step genome walking was done with primer pair AP1 and GSP3, and primer pair AP2 and GSG4. A DNA fragment of 449 bp from the Dra I genome walker library was identified as the upstream sequence of the PrAG1 promoter cloned from the first step genome walking based on the sequence comparison of overlapped region between them.

6) The 1105bp and 449 bp fragments were used in PCR mediated DNA splicing to synthesize one continuous 1458 bp promoter fragment of PrAG1. This was done as described. One primer was synthesized based on the 5'end sequence of 1105 bp promoter fragment: Primer PLi, 5' AGT TAC TTA ACA ATG CGC AAC CAA GGC 3'. Primer pair PLi and GSP2 was used in PCR to get the promoter fragment of 1105 bp, in which the AP2 primer sequence was removed. This 1105 bp fragment and 449 bp fragment was then added in one PCR tube as a template with the primer pair of AP2 and GAP2 to do the second round PCR to get the 1458 bp PCR fragment. The conditions of second round PCR were as follows: the first cycle at 95°C for 5 minutes, and 68°C for 10 min; the second cycle at 94°C for 30 seconds (DNA denaturing), DNA annealing at 60 C for 1 min, and DNA synthesis at 72 C for 2 minutes; this regime was cycled 30 times. This 1458 bp fragment was then cloned into pGEM-T easy vector (Promega) and subjected to DNA sequencing on both strands to confirm the DNA sequence and to make sure that no base changes occurred during the PCR process.

The sequence of the promoter is given in Figure 2.

7) DNA sequence analysis has indicated that compared to its orthologs from other plants, the PrAG1 showed that there were two possible positions for transcription initiation: at position 791 or 1326 from the promoter 5' end. It was found three typical TATA boxes in the PrAG1 promoter at the position of 280 to 286, 282 to 288,

1015 to1021. Based on the start codon position and short 5' untranslated region in the PrAG1 cDNA, the transcription initiation point is tentatively identified as position 1326 from 5' end of the promoter.

3. DNA and RNA Gel Blot Hybridizations:

Genomic DNA and RNA gel blots were made using standard techniques (Sambrook et al., 1989. Molecular Cloning : A Laboratory Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).

RNA: Total RNA was prepared from needle, vegetative shoot, stem, immature female cone and immature male cone samples as described above. Briefly, 20ug of total RNA was denatured in formaldehyde loading buffer and fractionated by denaturing agarose gel electrophoresis on a formaldehyde containing gel. The agarose gel was stained with ethidium bromide and a picture taken as control. The RNA was then transferred to a nylon membrane by the capillary blotting method. The RNA was immobilised on the membrane by UV cross-linking and was prehybridized at 65°C for 2 hours prior to hybridization in 0.5M Na-phosphate, pH 7.2, 7,5%SDS, 1mM EDTA, 100ug/Ml salmon sperm DNA. A DNA fragment of PrAG1 3' end region was labelled with 32P- dCTP (Decaprime II kit, Ambion, Austin, TX), and hybridised to the RNA blot overnight at 65°C. The blot was washed twice in 40mMNa-PO4, 1%SDS and 1mM EDTA for 30 minutes each at 65°C, and exposed to X-ray film with intensifying screens at –80°C.

DNA: Genomic DNA was prepared from needle tissue with CTAB method. Twenty μg genomic DNA was digested by Bam HI, Bgl II, Eco RI, Hind III and Xba I respectively. After agarose gel running, alkali blotting of DNA to Hybond N membranes was performed as described by the manufacturer (Amersham). The probe hybridisation and washing was as described for the RNA blotting analysis.

The results are shown in Figures 3 and 4.

4. RT-PCR:

Analysis was performed on total RNA isolated from needle, stem, vegetative shoot, immature female cone and immature male cone samples as described above. RNA

was reverse-transcribed with MMLV reverse-transcriptase (Gibco BRL) according to the manufacturer's instructions. PCR was performed with two primers: 5'PCR primer (5' TTGTGTACAAATCATGGG3') and 3'PCR primer (5' GTAAGCCCGTCACCCATC3'). Verification of the specificity of the PCR reactions was achieved through the use of controls that included amplification reaction with single primers, RNAse treatment of template, and no template. In those reactions in which no PCR product was detected, the quality of the RNA was tested by UV scanning, and agarose gel electrophoresis. ss-cDNA from the RT reaction was used as a template. The 50-ul reaction mixture contained 2.5 U Taq DNA polymease, 1X Polymerization Buffer (both from ClonTech Co.), 1mM MgCl2, 0.2mM dNTP and 0.25uM primers. The PCR was performed under following conditions: denaturation at 94°C for 30 s, annealing at 55°C for 1 min and extension at 72°C for 1 min for 30 cycles on Thermal Cycler 480 (Perkin-Elmer, Norwalk, CT, USA). The PCR products were subjected to electrophoresis in agarose gel, and hybridization as described above.

The results are shown in Figure 5.

Discussion

Northern blot hybridization and RT-PCR analysis showed that PrAG1 mRNA is accumulated specifically in the immature female cone and immature male cone; there is no expression detected in needle, stem, and vegetative shoot (Figures 3 and 5). This tissue distribution profile, when combined with the fact that PrAG1 contains a MADS box, verifies that PrAG1 is a reproductive gene in *Pinus radiata*.

Southern blot analysis showed that PrAG1 gene exists as a single copy in the genome of *Pinus radiata* (Figure 4).

INDUSTRIAL APPLICATION

In its primary aspect, the invention has application in modulating, and in particular reducing or eliminating reproductive capacity in plants of the *Pinus* genus. Such plants have utility in forestry.

The availability of reproductively null or sterile pine trees has the additional advantage that it will be possible to introduce further exogenous genetic material into those trees without the risk that the material will be passed on to other trees.

The invention also provides a new, reproductive-tissue-specific promoter. This promoter can be used in transforming plants of the *Pinus* genus as above, but also has wider application to other plants. In addition, the promoter can be used to drive expression of any gene which it is desirable to express in plant reproductive organs, including flowering time genes.

Those persons skilled in the art will appreciate that the specific description provided is exemplary only, and that modifications and variations may be made without departing from the scope of the invention.

RUSSELL McVEAGH WEST WALKER

per

ATTORNEYS FOR THE APPLICANT

```
TGT GTA CAA ATC ATG GGT CGT GGG AAG ATT GAG ATA AAG AGG ATT GAA AAT ACT  54
                M   G   R   G   K   I   E   I   K   R   I   E   N   T

ACG AAC CGA CAG GTC ACT TTC TGC AAG CGC CGA AAT GGT TTA TTA AAG AAG GCG  108
 T   N   R   Q   V   T   F   C   K   R   R   N   G   L   L   K   K   A

TAT GAA TTA TCA GTT CTT TGT GAT GCA GAA GTG GCC CTC ATC GTC TTC TCC AGC  162
 Y   E   L   S   V   L   C   D   A   E   V   A   L   I   V   F   S   S

AGA GGG AGA CTT TAT GAA TTT GCC AAC CAC AGC GTG AAG AGG ACG ATT GAG AGG  216
 R   G   R   L   Y   E   F   A   N   H   S   V   K   R   T   I   E   R

TAC AAG AAG ACT TGC GTT GAC AAC AAC CAC GGA GGG GCG ATA TCA GAG TCC AAT  270
 Y   K   K   T   C   V   D   N   N   H   G   G   A   I   S   E   S   N

TCT CAG TAT TGG CAA CAG GAG GCT GGT AAA CTC AGA CAA CAG ATT GAC ATT TTG  324
 S   Q   Y   W   Q   Q   E   A   G   K   L   R   Q   Q   I   D   I   L

CAA AAT GCA AAT AGG CAT TTG ATG GGT GAC GGG CTT ACA GCT TTG AAC ATT AAG  378
 Q   N   A   N   R   H   L   M   G   D   G   L   T   A   L   N   I   K

GAA CTC AAG CAA CTT GAG GTT CGA CTT GAA AAA GGA ATC AGC CGA GTG CGA TCC  432
 E   L   K   Q   L   E   V   R   L   E   K   G   I   S   R   V   R   S

AAA AAG AAC GAG ATG TTG CTT GAA GAG ATC GAC ATC ATG CAG AGA AGG GAA CAC  486
 K   K   N   E   M   L   L   E   E   I   D   I   M   Q   R   R   E   H

ATA CTT ATC CAG GAG AAT GAG ATT CTT CGC AGC AAG ATA GCC GAG TGT CAG AAT  540
 I   L   I   Q   E   N   E   I   L   R   S   K   I   A   E   C   Q   N

AGC CAC AAC ACG AAC ATG TTA TCA GCT CCG GAA TAT GAT GCA CTG CCC GCA TTC  594
 S   H   N   T   N   M   L   S   A   P   E   Y   D   A   L   P   A   F

GAC TCT CGA AAT TTC CTA CAT GCA AAT CTA ATC GAT GCG GCC CAT CAC TAT GCA  648
 D   S   R   N   F   L   H   A   N   L   I   D   A   A   H   H   Y   A

CAT CAG GAA CAA ACA ACG CTT CAG CTT GGC TGA ACG TTG AAG CGG TGG ACG CTT  702
 H   Q   E   Q   T   T   L   Q   L   G   *

AAA ACT CAA TCA AGG CAC CCG AAA AAT ATG CTA GTA ACC TTG AAT GAG ATT CAG  756

AGT CGA AAT ATT GCG AGG CAA GAG CAC AAT GGA AGA GAT AGC TCC TAG TAT GAA  810

TAT GGA TTT ATG ATA TTA ACA TAT GGT TTG TCA GCT TTA AAT ATA GCT GTT TGA  864

AAC AAA GAA TAC AAC ATA TTA GCT AGT ATT TTT TTG GCG CAT GTT ATC TTT CTG  918

TTG                                                                     921
```

Fig.1: The nucleotide sequence and its deduced amino acid sequence of PrAG1

```
AAACTCGACAGCAAATATGATTTAGATTATGACCTAGAAATAAGCATAGCATTAAAGCAT 60
ATACATAACAAGCGGTGATATACTCTGACTGCCACTGTACTTGAGGAAAGGTAGTGGACT 120
CTGCTCAGGTACATTAGTTTGGTAAGGTTGGCTTGGCTTCTGGGTAATATGAGAAGTAAA 180
GAAGTAAAAGGTATTTGACTCTAGTCAAGTACATTGGATTGCCTTTGTCGGGGCTTGGAT 240
GGCTTGGGTTCGTGTGAGAAGCCAACAATTTATAAGAAATATATAAAATAAAAAATAAAA 300
AAATTTAAGTGTTGGAAGTGAAAACGGTGGGGCAGAAATATACACAGAAGAGTACTTTAA 360
CAATGCGCAACCAAGGCAGATTCACAACTTGATTTCTGGACCTCGAATACGAGATAATGG 420
TGGTAAGAAATAAAGGAAGAGTGGAGTGCATTTGAAAATGAATGGAGAGCGCACAAAATG 480
GAGGACGAATAAATGAAATATAATGCAAGAGTGCATTTCCCTATTATTTCCAGAAATGTA 540
TATGTGGGGTCGGCATTCACATGGGCGTCGCATTCAGGGGGTGTCATAGCGGTCCTTTGA 600
TTGCAGTGTGGGAGTTGCAACATGTACCAACAAATCCATTCATCCCAAAACCTAAATTTA 660
TCCTCTCCATTACTATTACCTACACCTATACCTAGTAAATATGTCCTGCCTTGTAACTCC 720
TCCACTGCCTGCACACGTCTTAGTCAATCCATCTGCCTTCAAATAGGCATTATTTTGTTC 780
TTTCCCCTCCGACTGAAAGGCTATCGACCGACCGACCGCTCATCTTCTTCTTCTGCGCAA 840
TTTTTTCTGCTGGATCATCATCATTACCATCATCGCCATCCCCACCATCATCATCATGAT 900
GGTATCTCTATCTCTCCCTGGCAATCGATTGTAGAGGAAAGGAAGAGGGAAGGGGCATAT 960
GTATTGATCAACCTACCCGAAAAAACAATCTGATCAGCCCTGCTCAATCTTGCTTATAAA 1020
TCTCTTATCCACTGTTCAATCATTCAGGTTTCTTCCCACTTTCAAGCAAAGGCGCCCGGA 1080
TTGGCCGTGTTCTTAGATTTTCAGGTACTTAAATGGACAATATTCCCCACCTGAAGCCGT 1140
TCTGAAAAAGATTTGTTTGTAGAAACAAACGATTGTAATATTTGCTTAAGTTGAGCTTAA 1200
GGGGTTTGGTACCTAACTTGCCTTGTGGTTATTTGTTTCTCAGAACTCGGGCTGCGTCCA 1260
ACTGTAGGAACGAACCAGCACAAGGGGTTGCAGCTTTTGCTGTTGCTGTTGCGCCCATTG 1320
CTTTTGGACTGGTATTAGTAGTTGCAGCTTTGTTTTGCATACGCTGTGAGGATCTGTGCG 1380
                          M  G  R  G  K  I  E  I  K  R  I  E  N
CGGAAATTTTGTGTACAAATCATGGGTCGTGGGAAGATTGAGATAAAGAGGATTGAAAAT 1440
 T  T  N  R  Q  V
ACTACGAACCGACAGGTC                                              1458
```

Fig. 2: The promoter sequence of PrAG1

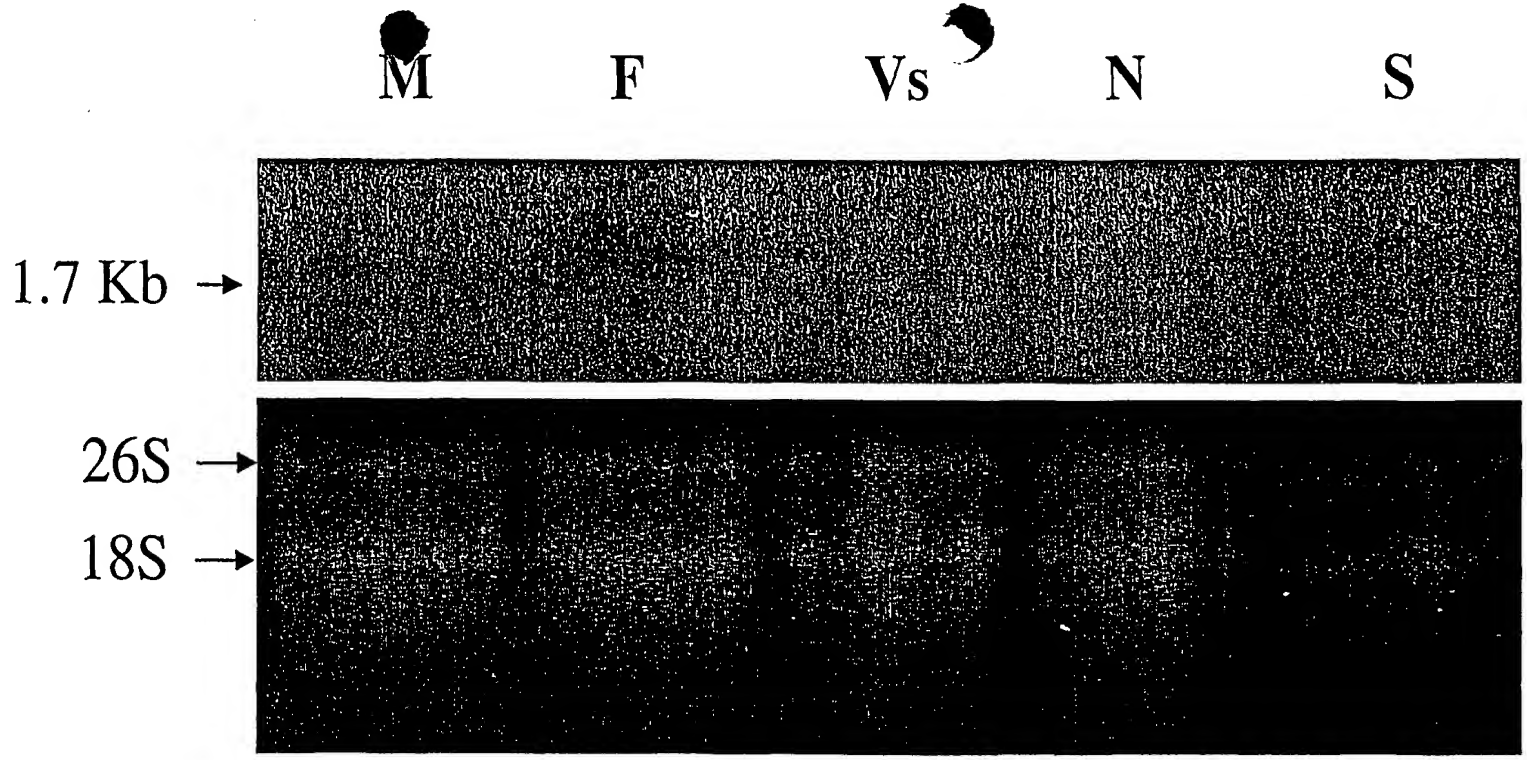

Fig. 3 RNA gel blot analysis of PrAG1 mRNA accumulation in radiata pine organs.

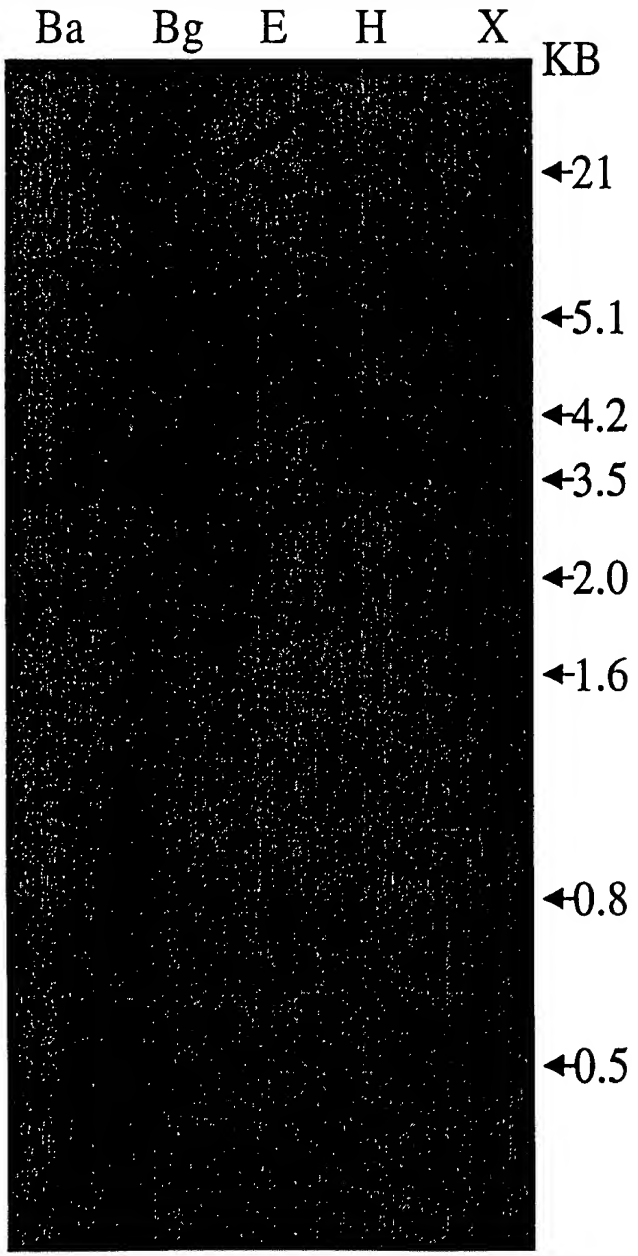

Fig. 4 DNA gel blot analysis of radiata pine genomic DNA hybridized with 3' terminal region of PrAG1 cDNA.

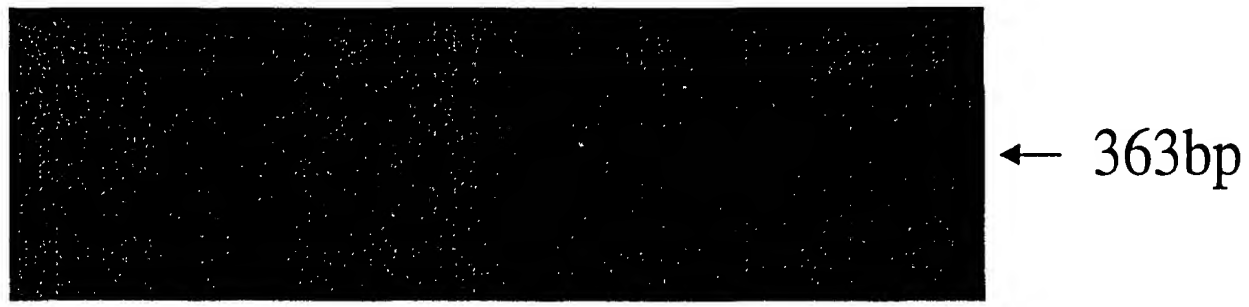

Fig. 5: Reverse transcription-polymerase chain reaction(RT-PCR) analysis showing reproductive organ-specific expression of PrAG1.